STILL PESSIMISTIC ABOUT FIRST-PERSON AUTHORITY:
A RESPONSE TO DOYLE AND WINOKUR

Wolfgang Barz

Abstract: This paper aims to support my (2018) skeptical position on the possibility of a correct and philosophically significant specification of first-person authority. For this purpose, I critically examine the proposals presented by Doyle (2021) and Winokur (2022) in response to my position and argue that while these proposals contain some ingenious ideas, they ultimately fall short of providing correct and philosophically significant specifications. Ultimately, the search for an adequate specification of first-person authority remains unresolved.

I. INTRODUCTION

In my paper, "Is There Anything to the Authority Thesis" (2018), I expressed pessimism regarding the feasibility of formulating a specification of first-person authority that is both true and philosophically interesting. This challenge was independently taken up by Casey Doyle (2021) and Benjamin Winokur (2022), both of whom presented intriguing suggestions aimed at refuting my pessimistic stance. Despite their ingenious proposals, which far surpass the subtlety found in much of the existing philosophical literature, Doyle and Winokur's arguments fail to achieve their goal. Ultimately, their proposed specifications of first-person authority confront the same dilemma I previously identified in my paper: they either turn out to be incorrect or lack philosophical interest. First, I provide a brief summary of the areas where Doyle, Winokur, and I concur, followed by a discussion of Doyle's and then Winokur's proposals.

II. COMMON GROUND: THE SEARCH FOR A SPECIFICATION OF FIRST-PERSON AUTHORITY

Doyle, Winokur, and I agree on seeking a *satisfactory specification of the phenomenon of first-person authority*. Note three things about it. First, "first-person authority" refers to a characteristic that applies to sentences like "I want to eat cookies," "I believe that global warming is an urgent matter," and "I'm in pain" (as per Winokur's 2022 examples). Therefore, it is important not to conflate first-person authority with

neighboring ideas like epistemic privilege or peculiar method.[1] While epistemic privilege and peculiar method are ideas pertaining to the beliefs we form about our mental life and the ways we achieve knowledge about it, first-person authority is "a property of utterances or speech acts" (Doyle 2021: 364); it is "a kind of epistemic status that others cede, or can justifiably cede" to self-ascriptions (Winokur 2022: 116).

Second, the phrase "specification of the phenomenon" highlights that the focus is not on *explaining* first-person authority, but on *describing* the phenomenon that such an explanation would account for. Our question is not, "*Why* do listeners grant special authority to self-ascriptions?" but rather "*What is* this special authority that listeners grant to self-ascriptions?" The reason for framing the question this way is that many descriptions of the phenomenon of first-person authority in the literature are, as Doyle notes, "muddled… or vague" (Doyle 2021: 363). If the explanandum is unclear, it does not bode well for the quality of the explanation. Therefore, the issue is a sort of applied meta-philosophy. We step back, take a broader view of the philosophical debate on first-person authority, and ask: what exactly is the phenomenon that philosophers are trying to explain, and what makes it so philosophically puzzling?

Third, the phrase "satisfactory specification" suggests that not just any arbitrary specification of the phenomenon of first-person authority will suffice. Rather, any specification must meet at least three criteria, which Doyle helpfully labels Prevalence, Asymmetry, and Wonder (see Barz 2018: 127; Doyle 2021: 365; Winokur 2022: 116). Prevalence excludes specifications that do not apply to self-ascriptions, or only apply to a small number of them. Asymmetry excludes specifications that also apply to other-ascriptions, i.e., to ascriptions of mental states from a third-person perspective. Thus, Asymmetry aims to account for the common assertion that first-person authority is a unique property of self-ascriptions that is not found elsewhere. Finally, Wonder eliminates specifications that describe the phenomenon of first-person authority in such a way that the resources of philosophy are unnecessary to explain it. This condition aims to account for the fact that philosophers generally consider first-person authority to be within their area of expertise. If first-person authority were a phenomenon that could be explained by ordinary psychological research, it would be difficult to understand why so many philosophers write books and essays on the subject.

Although Winokur does offer some criticisms of Prevalence and Wonder (2022: 117-118), the three of us agree on the importance of all three criteria. The crux of our

disagreement lies elsewhere: Doyle and Winokur contend that there exist specifications that satisfy all three criteria, whereas I disagree.

III. IS A SPECIAL EXPLANATION NEEDED?

Doyle proposes the following specification of the phenomenon of first-person authority:

> When a subject self-ascribes a current conscious mental state in speech, there is a presumption that what she says is true. In order to be epistemically justified in defeating the presumption one must possess *both* evidence that what she says is false *and* evidence that entitles one to an explanation of how she has been led astray. (Doyle 2021: 367)

Doyle understands this proposal as a constructive development of a proposal I had criticized as inadequate in my (2018) paper. To understand the point of Doyle's proposal, it may be helpful, then, to briefly review my critique. The claim I had criticized was that the hearer — provided they lack any good reason to doubt what the speaker says — is justified in believing a self-ascription to be true even if they are not in possession of any supporting evidence ("Presumption of Truth$_1$"). The problem with Presumption of Truth$_1$ (and Doyle and Winokur agree) is that it violates either Prevalence or Asymmetry. It violates Asymmetry if one subscribes to a liberal view on testimony, according to which the possession of supporting evidence is not required to be justified in believing the testimony of another person to be true (given the absence of reasons to doubt what the speaker says).[2] Viewed from this liberal perspective on testimony, not only self-ascriptions but also *other-ascriptions* enjoy the property specified in "Presumption of Truth$_1$": they, too, may be justifiably believed to be true—even in the absence of supporting evidence—as long as there is no reason to doubt what the speaker says.

Conversely, if we subscribe to a conservative view of testimony, Presumption of Truth$_1$ violates Prevalence. According to conservatives regarding testimony, it is not enough to lack any reason to doubt what the speaker says. Rather, to be justified in believing what another person says to be true we must also possess some supporting evidence. If conservatism is true, then self-ascriptions lack the property specified in "Presumption of Truth$_1$": even if there is no reason to doubt what the speaker says, self-

ascriptions could only be justifiably believed to be true as long as the hearer possesses some supporting evidence. In short, it does not matter whether one leans towards the liberal or conservative view on testimony—either way, Presumption of Truth$_1$ proves problematic.[3]

Doyle's proposal is meant to avoid this dilemma by centering on the idea of a "demand for a special explanation." Doyle argues that this demand is uniquely applicable to self-ascriptions, but not to other-ascriptions. To illustrate this point, Doyle provides the following example:

Suppose I tell you that our mutual friend, Tom, is very upset that the Houston Astros lost last night. […] But Tom is your friend, too, and you know a thing or two about him. You know that Tom is not a baseball fan. So you have good reason to believe that Tom does not feel any particular way about the Astros's fate. It seems plausible that this is all you need to defeat the reason provided by my testimony. Knowing what you do about Tom is a good enough reason not to take me at my word. This is so even though you may well lack a story about why I came to a different conclusion about Tom. You have got a good reason to believe that what I say is false and so a good reason to believe that I am wrong. You do not need to explain why I am wrong to defeat my testimony.

But things are different when it comes to self-ascriptions. Suppose I tell you that I want a cup of coffee. As my friend, you know that I am not a coffee drinker. So it is surprising to hear me say this. You have some evidence for thinking that I am wrong, that I do not really want coffee. Still, this is not enough to defeat my testimony. You should still take me at my word, unless you have some story in hand about why I have been led astray. Any number of such explanations might be available. Perhaps I am being insincere and am merely trying to fit in with a peer group. Perhaps I am confused, drunk, self-deceived, or otherwise irrational. Different explanations will apply to different cases. But, arguably, in order to justifiably refrain from taking me at my word you need to be entitled to some such explanation. (Doyle 2021: 367)

I have some queries about Doyle's proposal. He explicitly mentions that the hearer does not need to *know* why the speaker is wrong. According to Doyle (2021: 368), the

explanation of why the speaker is wrong *may be incorrect*. This provision creates a gap for counter-examples. Let's revisit the coffee-case again: I ask you for a cup of coffee, but you are surprised because you always believed that I don't like coffee. In the past, I have refused offers to drink coffee saying, "No thanks, I don't like coffee." But now, I want coffee, and the question is: should you take me at my word, or should you assume that I am mistaken? Let us elaborate on this scenario further. Suppose that while expressing my desire for coffee, you notice a slight irregularity in my blinking. According to your belief system (which you have largely constructed based on consuming conspiracy theory videos on Youtube), this observation suggests the following explanation: I have been abducted by aliens who implanted a mechanism in my brain that systematically distorts my judgments about what I want. The conditions of Doyle's proposal are satisfied: you have evidence that what I say is false, and evidence that entitles you to an explanation of how I have been misled. Thus, according to Doyle's proposal, you are epistemically justified in rejecting the presumption that what I say ("I want a cup of coffee") is true. However, this strikes me as highly counterintuitive. Rejecting the presumption on such a preposterous basis would be epistemically irresponsible.

An obvious objection to my argument is that it relies on an unfair interpretation of the concept of "evidence" regarding Doyle's proposal. What does possessing evidence actually entitle someone to? Is it having beliefs that one *believes* are supported by the evidence, or having beliefs that *are actually* supported by the evidence? My argument is based on the first, subjective interpretation of "evidence." However, if we consider the objective interpretation of "evidence," it could be more favorable to Doyle's proposal. Under this interpretation, you may believe that the observed irregularity in my blinking supports the conspiracy theory explanation, but in fact, it does not support it. Therefore, if we adopt this objective reading of "evidence," observing the irregularity in my blinking does not entitle you to an explanation of how I have been misled. Consequently, you are not justified in defeating the presumption that what I say ("I want a cup of coffee") is true. This resolves the problem.

However, even if we adopt the objective reading of "evidence," Doyle's proposal still faces difficulties. To highlight these difficulties, it is useful to examine a variant of Schwitzgebel's (2012) case of Ralph, which Doyle also discusses. Ralph, a distinguished philosophy professor, behaves in a sexist manner, treating women as if they are less intelligent than men, systematically discriminating against them in job

allocation, and demanding stronger proof of their competence than he would for men. Therefore, when Ralph tells us, "I believe that men and women are equally intelligent," it is evident to us (and to everyone who is familiar with Ralph) that what he says is simply wrong.[4] Let us further assume that Ralph provides us with little or no indication as to *why* he would make such an obviously untrue statement about himself. So it does not seem feasible to determine precisely what accounts for Ralph's utterance of such an entirely false sentence. Is he insincere, meaning that he disbelieves what he says, but verbally pretends to believe it in order to avoid losing the appreciation of his academic peer group? Or is he sincere, meaning that he believes what he says, but systematically disregards any contrary evidence in order to avoid damaging his highly moral self-image? Let us assume that each of these explanations is equally supported by the available evidence. This implies that the evidence we possess does not entitle us to any specific explanation. Therefore, according to Doyle's proposal, we are not epistemically justified in defeating the presumption that what Ralph says ("I believe that men and women are equally intelligent") is true. We must take Ralph at his word, even though we possess overwhelming evidence to the contrary. This is counterintuitive.

Again, one might object that my argument is based on an unfavorable reading of Doyle's proposal. Doyle's crucial claim is:

> one must possess […] evidence that entitles one to an explanation of how [the speaker] has been led astray. (Doyle 2021: 367)

There are two ways to understand this: first, the evidence must entitle the subject to a *specific* explanation, and, second, the evidence must entitle the subject to *some* explanation. The first reading might be paraphrased thus:

> $\exists x$ (x is an explanation of how the speaker has been led astray & one must possess evidence that entitles one to believe that x is true)

The second reading thus:

> one must possess evidence that entitles one to believe that $\exists x$ (x is an explanation of how the speaker has been led astray)

The problem I raised for Doyle, one might say, exists only if we adopt the first reading. So we should rather adopt the second reading.

However, the second reading is evidently incompatible with Doyle's intended meaning, as Doyle explicitly states that we must possess evidence that entitles us to a *specific* explanation (see Doyle 2021: 368). But the real difficulty with Doyle's proposal lies elsewhere. The problem, which Doyle anticipates (2021: 372, fn.16), is that having evidence that what the speaker says is false *is equivalent to* having evidence that there is some explanation for how the speaker has been led astray. Everything that happens occurs for a reason or cause. Thus, there is always an explanation for anything that occurs. Therefore, as soon as one has evidence that someone's self-ascription is incorrect, one is justified in believing that there is some explanation for the speaker's mistake. This implies that Doyle's demand for evidence that entitles one to some explanation of how the speaker has been led astray is redundant. The demand is automatically met by possessing evidence that what the speaker says is false. Consequently, Doyle's proposal reduces to the following:

> When a subject self-ascribes a current conscious mental state in speech, there is a presumption that what she says is true. In order to be epistemically justified in defeating the presumption one must possess evidence that what she says is false.

However, this specification violates Asymmetry because exactly the same is true of other-ascriptions, at least if we take a liberal view of testimony. On the other hand, if one prefers a conservative view of testimony, then the specification violates Prevalence. For in this case hearers are not entitled to presume a self-ascription's truth as long as they lack positive evidence.[5] As a result, Doyle's proposal is confronted with the same dilemma that Presumption of Truth$_1$ previously encountered.

IV. DEGREES OF JUSTIFICATION

Winokur (2022) suggests three different specifications of first-person authority. His first proposal revolves around the idea of a presumption of truth, while the other two proposals make use of the idea of groundlessness. In this section, I discuss Winokur's first proposal. I discuss the other two proposals in the next sections. Winokur's first proposal reads as follows:

> If (1) a person *S* tells another person *H* at time *t* that she (*S*) is currently in mental state ɸ and (2) *H* lacks any good reason to doubt what *S* says, then (3) *H* is more

7

justified at t in believing that S is currently in ɸ than H would be in believing an ascription of ɸ to S by some other person, P, even if H lacks any reason to doubt P, and even if H possesses no positive evidence that S is currently in ɸ (aside from the fact that S or P tells him so). (Winokur 2022: 121).

To simplify matters for the following discussion, let us compare two situations:

Situation 1. Susan tells Herold, that she, Susan, has a toothache. Apart from the fact that Susan tells him so, Herold has neither evidence to the effect that Susan has a toothache (positive evidence) nor evidence to the effect that Susan does not have a toothache (negative evidence). Thus, Herold is in a kind of evidential vacuum: the only piece of evidence he possesses is Susan's self-ascription ("I have a toothache").

Situation 2. Peggy tells Herold that Susan has a toothache. Again, Herold is in a kind of evidential vacuum: he lacks both positive and negative evidence for what Peggy tells him is true. The only piece of evidence he possesses is Peggy's other-ascription ("Susan has a toothache").

Like Doyle, Winokur views his proposal as a constructive development of Presumption of Truth$_1$. Presumption of Truth$_1$ suggests that in Situation 1, Herold is justified in believing that Susan has a toothache, while in Situation 2, he is not. This strikes all three of us — Doyle, Winokur, and me — as odd since we have the liberal intuition that even in Situation 2, Herold is justified in believing that Susan has a toothache. However, simply conceding that Herold is *just as* justified in Situation 2 as in Situation 1 would violate Asymmetry. To address this issue, Winokur emphasizes that justification comes in degrees. According to Winokur, in Situation 1, Herold is *more* justified in believing than he is in Situation 2. However, this does not imply that in Situation 2, Herold has no justification at all. Rather, in Situation 2, he has some justification to believe that Susan has a toothache, albeit to a *lesser degree* than in Situation 1.

The problem with Winokur's proposal is that, given an evidential vacuum, self-ascriptions do not provide the hearer with a higher degree of justification than other-ascriptions. Though Winokur already outlined my argument for this claim (see Winokur 2022: 122), I would like to rephrase it using the example of Herold, Susan, and Peggy to clarify.

To better understand the problem with Winokur's proposal, let's examine how much additional evidence would be required in each situation for Herold's belief that Susan has a toothache to reach the highest level of justification. By determining the necessary amount of evidence, we can calculate Herold's initial degree of justification by subtracting the degree of justification provided by the supplementary evidence from the maximum level of justification.

Let us examine Situation 1 first, where Herold only has Susan's self-ascription, "I have a toothache," as evidence. Even though this self-ascription justifies Herold to some extent in believing that Susan has a toothache, the highest level of justification has not yet been achieved. Generally, there are three types of evidence that Herold would require to increase his justification to the maximum level possible: non-linguistic behavioral evidence, bodily evidence, and neurophysiological evidence. For instance, if Herold noticed that Susan was groaning, avoiding chewing, or putting her hand to her cheek (non-linguistic behavioral evidence); and if he observed that Susan had some tooth decay (bodily evidence); and if he identified through Magnetic Resonance Imaging that Susan's C-fibers were firing (neurophysiological evidence), then Herold would probably have attained the highest degree of justification.

Let us turn to Situation 2, where Herold only has Peggy's other-ascription, "Susan has a toothache," as evidence. As with Situation 1, the maximum degree of justification has not been reached. Now, consider the following scenario: while Peggy tells Herold that Susan has a toothache, Susan is standing next to Peggy, allowing Herold to observe her groaning, avoiding chewing, and putting her hand to her cheek. Upon looking into Susan's mouth, he also sees that she has tooth decay. Furthermore, since Susan happens to be hooked up to an MRI scanner, Herold receives confirmation that her C-fibers are firing. It is highly plausible that this additional evidence would raise Herold's degree of justification to its maximum.

The issue with Winokur's proposal is now apparent: since the same amount and type of supplementary evidence is needed to reach the highest degree of justification in both situations, Herold's initial degree of justification in Situation 1 is equal to his initial degree of justification in Situation 2.

Winokur presents an intriguing response to the previous argument by questioning the assumption that the same amount and types of evidence have the same justificatory force in both situations. According to the "transmission" view of testimonial justification, for example, it is possible for the same amount and types of evidence to

have different justificatory force in different situations. Therefore, although Herold may need the same amount and types of additional evidence in both Situation 1 and Situation 2 to achieve the maximum level of justification, it is possible that the degree by which the same amount and types of evidence increases justification is lower in Situation 1 than in Situation 2. As a result, the degree of Herold's initial justification in Situation 1 could be higher than in Situation 2 (Winokur 2022: 122-123).

But it's unclear that the problem for Winokur's proposal can be eliminated in this way. Let's assume, for the sake of argument, that there is some truth to the claim that the justificatory force of neurophysiological, bodily, and behavioral evidence is worth less in Situation 1 than in Situation 2. If this is true for *positive* evidence, then it should also hold for *negative* evidence. Let's say that all evidence points to Susan *not* having a toothache: she's laughing, biting into an apple, her teeth are fine, and her MRI shows that her C-fibers are at rest. According to Winokur, this evidence would have different justificatory force in the two situations, with it being worth less in Situation 1 than in Situation 2. This leads to the conclusion that, despite overwhelming evidence to the contrary, Herold should still trust Susan's self-ascription a little more than Peggy's other-ascription. However, this claim is counterintuitive. The negative evidence has the *same* justificatory force in both situations: it decreases the initial justification in both situations to the same degree, namely, to zero. If neurophysiological, bodily, and behavioral evidence have the same justificatory force when it comes to *decreasing* justification, why should they have different justificatory force when it comes to *increasing* justification?[6]

V. GROUNDLESSNESS 1

Winokur proposes two more specifications of the phenomenon of first-person authority. The first reads thus:

> If (1) a person *S* tells another person *H* at time *t* that she (*S*) is currently in some mental state ɸ and (2) *H* lacks any good reason to doubt what *S* says, then *H*'s justification for believing that *S* is in ɸ is not undermined if there are no epistemic reasons *S* can reasonably provide for her self-ascription if requested to do so by *H* (such that any request for them from *H* is inappropriate) (Winokur 2022: 126).[7]

Applying Winokur's proposal to the example of Herold, Peggy and Susan, we get the following: in Situation 1, Herold's initial justification is not undercut if Susan is not able to provide an epistemic reason (as to why she believes that she has a toothache) if requested to do so by Herold—whereas, in Situation 2, Herold's initial justification *is* undercut if Peggy is not able to provide an epistemic reason (as to why she believes that Susan has a toothache) if requested to do so by Herold.

However, it is doubtful that Herold's justification in Situation 1 is not undermined if Susan is unable to provide epistemic reasons. Let's say Herold asks Susan, "How do you know you have a toothache?" and Susan responds with a shrug, "I don't know how I know it, the thought I have a toothache just randomly popped into my head, I don't believe it for any reason." This answer would be unsatisfactory. Consequently, Herold's initial justification for believing that Susan has a toothache would rapidly decrease at that point.

Now, consider Herold asking Susan, "How do you know you have a toothache?" and Susan replying, "What a silly question, Herold! I know it because I feel the pain in my tooth!" This is a completely appropriate response, exactly the response we would expect. Therefore, Winokur's proposal violates Prevalence: the specified property (namely, being such that the hearer's initial justification is not undermined if the speaker is unable to provide epistemic reasons) does not apply to self-ascriptions.[8]

Winokur anticipates this objection. He admits that *phenomenal* self-ascriptions (such as "I have a toothache") might reasonably be viewed as based on epistemic reasons that the subject in question could provide on request. However, according to Winokur, such a view is not very plausible when it comes to self-ascriptions of *propositional attitudes*. He backs up this claim with the following example:

> *H* enters a party and sees *S* leaving. *H* asks *S* why she is leaving and *S* tells *H* that she believes that staying too long at parties brings bad juju. [The proposal at hand] claims that there is nothing *S* can be reasonably expected to say in response to the question "how do you know you believe that?" In this sense, the question is inappropriate. (Winokur 2022: 127)

It may be true, then, concludes Winokur, that his proposal does not apply to self-ascriptions of phenomenal states. However, it applies to self-ascriptions of propositional

attitudes. Since self-ascriptions of propositional attitudes form a sufficiently large set of self-ascriptions, Prevalence is not violated.

I do not find this move particularly convincing, as I argued at length in several other places that, in general, we gain first-personal self-knowledge by applying the so-called transparency method. In my view, knowledge of one's own belief that p, for example, can be rationally reconstructed as the outcome of an inference beginning with the premise "p." Similarly for other attitudes.[9] Thus, according to my approach, there are epistemic reasons on which self-knowledge of propositional attitudes is based, and which the subject in question could provide upon request.

However, one need not be a proponent of the transparency approach to acknowledge that self-knowledge regarding one's propositional attitudes is based on epistemic reasons. Even Brie Gertler, who is a vocal critic of the transparency approach, would concur with this claim. In the current context, one of Gertler's examples is particularly interesting, as it revolves around a superstitious belief, similar to Winokur's example cited above. Specifically, Gertler's example features a person named Nick, who was "raised to believe that spilling salt will bring bad luck, which can be averted only by immediately dropping a pinch of salt over one's shoulder" (Gertler 2011: 135). Nick, now grown up, knows that this belief is pure superstition, yet he continues to have it: "whenever he spills salt he has an immediate sense of doom, and finds himself compelled to drop a few grains over his shoulder" (Gertler 2011: 135). Gertler denies that Nick can acquire self-knowledge about his belief using the transparency method. According to Gertler, the reason for this is that the transparency method requires one to consider whether it would be rational to believe that spilling salt brings bad luck. If the answer is "yes," then one is justified in believing that one has the belief in question. The problem with Nick is that if he were to use the transparency method, he would conclude that it is *not* rational to have the belief in question (since he knows that it is pure superstition). Thus, according to the transparency approach, he would not be justified in believing that he believes that spilling salt brings bad luck. Thus, Gertler concludes, the transparency approach cannot explain how Nick knows that he has the belief that spilling salt brings bad luck. Instead, Gertler proposes to reconstruct Nick's self-knowledge as the result of an "imaginative exercise":

Nick […] pictures some salt falling from a shaker in his hand. As he visualizes the grains dropping to the floor, he is full of foreboding, and feels a strong urge to

pour salt over his shoulder. He concludes "I guess I still believe that spilling salt brings bad luck." (Gertler 2011: 135–136)

Now let's go back to Winokur's example of the person who, when asked why they are leaving the party, responds with "I believe that staying too long at parties brings bad juju." Building on Gertler's ideas, one might say that there are several reasonable answers that the speaker might give in response to the question, "How do you know you believe that?" Consider the following response, for instance:

Well, that's a difficult question. Let me think for a moment ... I think I know it because when I imagine myself staying too long at a party, I tend to get an uneasy feeling in my stomach. It may sound silly, but this feeling is accompanied involuntarily by scary thoughts about my future: of me being hit by a car, of my boss firing me, or of a natural disaster striking us all. Now, I'm not a psychologist, but it seems to me that my mental reactions to the mere idea of staying too long at a party clearly indicate that I have the belief that staying too long at parties brings bad juju. What do you think?

I find this answer to be perfectly reasonable. As such, I have doubts about Winokur's claim that self-ascriptions about propositional attitudes are groundless.[10] From an epistemic perspective,[11] it is always legitimate to ask for reasons behind a self-ascription, whether it pertains to a phenomenal experience or propositional attitude. Asking for reasons is not inappropriate in either case. Of course, this is not to say that Winokur's proposal is simply wrong, but it is dialectically weak: it relies on the idea of groundlessness, which is itself subject to controversy and debate.

VI. GROUNDLESSNESS 2

Winokur's third proposal for specifying the phenomenon of first-person authority reads thus:

If (1) a person S tells another person H at time t that she (S) is currently in some mental state ϕ and (2) H lacks any good reason to doubt what S says, then, given normal circumstances, H's willingness to question whether S really is in ϕ is

rather low, even though there is no epistemic reason *S* can reasonably provide for her self-ascription if requested to do so by *H* (such that any request from *H* is inappropriate). (Winokur 2022: 127)[12]

This proposal is structurally similar to the previous one, though it does not aim at the hearer's *justification*, but at the hearer's *willingness to question* what the speaker is saying. As the concept of groundlessness is being used again, my previous concerns still stand. But even if we put those aside for now and assume that self-ascriptions are groundless, I still see difficulties for Winokur's proposal.

The first difficulty is that it is quite natural and expected that the hearer's willingness to question a self-ascription will be low. After all, if we assume that there are no reasons to provide for a self-ascription, why would we request them? Thus, given that self-ascriptions are groundless (as we assume for the sake of argument), Winokur's proposal offers a simple explanation for why our willingness to question self-ascriptions is low: any request for reasons would be considered inappropriate. As we generally seek to avoid inappropriate behavior, we tend to shy away from questioning what the speaker says. Thus, the two facts that Winokur sees as being in tension—that (i) our willingness to question self-ascriptions is low and (ii) self-ascriptions are groundless—actually fit together quite well. There is no real tension between them: fact (ii) explains fact (i) in a straightforward manner.

The second difficulty is that Winokur's proposal specifies the phenomenon of first-person authority in terms of a *willingness to question* what the speaker says. The willingness to do something—whether it's questioning self-ascriptions, helping those in need, voting in the presidential election, or what have you—is a behavioral disposition. The question of why we have (or do not have) a particular behavioral disposition is a question for empirical psychology to answer, not philosophy. Thus, Winokur's proposal seems to violate Wonder.

Winokur, of course, anticipates this objection. He admits that, in light of his proposal, first-person authority is "a fact about the psychological dispositions of rational agents, and so counts as an empirical-psychological phenomenon" (Winokur 2022: 128). Nevertheless, he insists that there is a genuine philosophical problem: "epistemologists should puzzle over the fact, when dealing with self-ascriptions, that *H*'s belief-forming practices diverge from a basic epistemic norm according to which

we ought to place greater trust in the person who can epistemically support her claims" (Winokur 2022: 128).

The trouble with this reply is that hearers' reluctance to question self-ascriptions does not constitute a deviation from any epistemic norm, *given that self-ascriptions are groundless*. In order to see the point, take a look at the epistemic norm Winokur brings into play:

(TRUST) You ought to place greater trust in the person who can epistemically support her claims!

The purpose of TRUST is to guide us in cases where a person says something that can, in principle, be supported by reasons. Suppose you are talking to a friend on the phone and she tells you that there has just been a traffic accident at the intersection where she lives. When you ask her how she knows, she replies: "There was just a squeal and a terrible bang shortly after. It sounded as if two cars had crashed. That happens quite often here. I know that sound." TRUST states that, as far as the alleged traffic accident is concerned, you should place greater trust in your friend than in a person who simply replied: "I've no idea, it just popped into my head."

However, when the speaker's utterance belongs to a category that cannot be supported by any reasons, things are different. Imagine that you are attending a business dinner and the person sitting next to you asks you to pass her the salt, saying, "Could you pass the salt, please?" In this situation, if you ask, "How do you know?" the person will either be confused or think it's a bad joke. In either case, she will not be able to give a reasonable answer to your question. It would be odd to say that TRUST obliges us not to accept what is said ("Could you pass the salt, please?") because she speaker cannot cite any reasons. Rather, the appropriate response is to say that TRUST is simply not applicable in this case. The boundary conditions for its application are not met: the speaker says something that could not, not even in principle, be substantiated by reasons.[13]

Now, *given that self-ascriptions are groundless*, exactly the same applies to situations in which someone utters a self-ascription: since the speaker does not say anything that could be substantiated by reasons, the boundary conditions for TRUST's application are not met. Consequently, TRUST does not apply. If TRUST does not apply, Winokur's idea that hearers are *deviating* from TRUST if they take self-ascriptions at

face value (or are not very willing to question them) is simply misplaced—just as it would be misplaced to say that you are deviating from TRUST if you grant the person's request to pass the salt. If hearers do not deviate from TRUST if they take self-ascriptions at face value, there is nothing for epistemologists to puzzle over. Thus, Winokur's proposal violates Wonder.

VII. SUMMARY

I critically evaluated four proposals presented by Doyle (2021) and Winokur (2022) in response to my (2018) argument on first-person authority. However, none of these proposals meets the necessary requirements for an adequate specification. Doyle's proposal that the hearer of a self-ascription must have evidence of how the speaker has been led astray to defeat the initial presumption of truth is too strong on one interpretation and too weak on another. Winokur's first proposal, which involves different degrees of justification, is problematic because hearers of self-ascriptions are justified to the same degree as hearers of other-ascriptions, all things being equal. Both Winokur's second and third proposals are flawed since they assume that self-ascriptions are groundless, which is highly controversial. Even if we accept that self-ascriptions are groundless, Winokur's third proposal fails to explain why first-person authority is philosophically puzzling. As a result, I conclude that an adequate specification of first-person authority has not been achieved.

Of course, none of the preceding reflections implies that finding a specification that satisfies all three constraints accepted by Doyle, Winokur, and me, i.e., Prevalence, Asymmetry, and Wonder, will be impossible.[14] However, given this paper's discussion, it will be quite challenging to find one. Thus, the more promising path for my opponents might be to challenge one or more of the above constraints. But which? Should we abandon Prevalence and argue that first-person authority pertains only to a vanishingly small number of self-ascriptions? Should we abandon Asymmetry and posit that first-person authority extends to other-ascriptions besides self-ascriptions? Or should we abandon Wonder and claim that first-person authority poses no philosophical puzzle and can be explained solely by empirical sciences? In my view, the answer to each of these questions is unequivocally "no." Therefore, I remain pessimistic about the prospect of ever finding a satisfactory specification of first-person authority.[15]

ENDNOTES

1 The idea of epistemic privilege is that "beliefs about one's mental states acquired through the usual route are more likely to amount to knowledge than beliefs about others' mental states" (Byrne 2018: 5). The idea of peculiar method, on the other hand, is that "one can come to know about one's mental life in a way that is available to no one else" (Byrne 2018: 8).

2 I use the labels "liberal" and "conservative" in the sense of Green 2008, who explains this usage as follows: "Liberals are less demanding on testimonially-based justification and allow testimonially-based beliefs to count as justified, or as knowledge, more liberally; conservatives are more demanding and dispense testimonially-based epistemic honors more conservatively." Doyle cites Burge's 1993 Acceptance Principle as an example of a liberal view regarding testimonially-based justification. An example of a conservative view is Lackey 2008.

3 An anonymous reviewer pointed out that some views of self-knowledge, such as expressivists in the tradition of the late Wittgenstein, would actually deny that uttering a self-ascription involves engaging in testimonial practices. Thompson 2012 attributes this view to Anscombe, stating: "In the case where I tell you that I'm in pain or that I think that something is wrong with my liver, you don't have it secondhand by simple knowledge transfer; nor, on the other hand, is your knowledge of my pain a mere matter of observation exactly. When it comes to the narrowly psychical features of the speaking animals, the humans—well, there is, among them, a way of coming to know certain propositions about one another from one another that is sui generis, or anyway, in need of another account: *it is neither testamentary in any ordinary sense*, nor observational, nor is it derived by inference" (199, emphasis added). It appears, then, that an expressivist position could potentially escape the dilemma presented. I don't think so, but due to space constraints, suffice it to mention that both Doyle (2015: 84, ft 64), and Winokur (2022: 129, fn. 11), are also skeptical about whether adopting an expressivist position could effectively avoid the dilemma.

[4] The case I describe here differs significantly from Schwitzgebel's 2012 original case. In mine, the evidence strongly indicates that Ralph does not hold the belief he claims to hold. As a result, it is readily apparent to all observers that Ralph's self-ascription ("I believe that men and women are equally intelligent") is false. Schwitzgebel's case is more intricate, as he portrays Ralph as being "prepared to argue coherently, authentically, and vehemently for the equality of intelligence" (2012: 191). Thus, Schwitzgebel's Ralph case is widely regarded as a quintessential instance of "implicit bias," a phenomenon that has sparked heated debates over its explanation. The crux of this discussion pits those who believe that implicit bias can be best explained by positing conflicting propositional attitudes (as traditionally understood) against those who seek alternative explanations without relying on propositional attitudes. Schwitzgebel aligns with the latter approach and endeavors to elucidate instances of implicit bias using the notion of an "in-between state." (For a comprehensive overview of this debate, refer to Brownstein 2019. For critical discussions of in-between states in the context of clinical delusions, see Bortolotti, 2012: 40-48.) However, note that the question of explicating implicit bias is irrelevant to the case I present, in which there is unequivocal certainty that Ralph does not believe in the intellectual equality of genders. His self-attributions are mere lip service. Consequently, the question of implicit bias doesn't arise. The focal point is simply that the observer cannot discern *why* Ralph makes a statement about himself that glaringly contradicts the evidence.

[5] Doyle's formulation does not explicitly state that the presumption exists even when the hearer lacks positive evidence. Nevertheless, based on Doyle's argumentation, it is apparent that he intends this understanding of his formulation.

[6] An anonymous reviewer raised a concern about the case of Susan not having a toothache, stating it's underexplained as it remains unclear whether she is lying or being self-deceived. Depending on Susan's situation, the argument could lead to different conclusions. However, I disagree with this assessment. The question of *why* Susan says something untrue is irrelevant to the justificatory force of the evidence at hand. To illustrate this point, consider two cases: (i) Susan is lying when she says "I have a toothache," and (ii) Susan is self-deceived when she says "I have a toothache." I fail to see why the justificatory force of negative evidence (Susan laughs, bites into an apple, her teeth are fine, and her C-fibers are at rest) should vary in both cases.

[7] I here ignore the so-called "complementary clause," which states that the equivalent does not apply to other-ascriptions, i.e., if there are no epistemic reasons that the utterer of an other-ascription can give in response to the listener's request, the justification *is* undermined.

[8] The comparison between the two dialogues reveals another point: the inappropriateness of requests for epistemic reasons for self-ascriptions may not be related to the absence of such reasons. In the second dialogue, Susan clearly perceives Herold's question as inappropriate ("What a silly question, Herold!"), but the reason is not that she has no epistemic reason for her self-ascription. In my view, Winokur's proposal thus mixes two ideas that should be better kept apart: the groundlessness of self-ascriptions with the inappropriateness of requests for epistemic reasons for self-ascriptions. (On the subject of inappropriateness, see also endnote 11.)
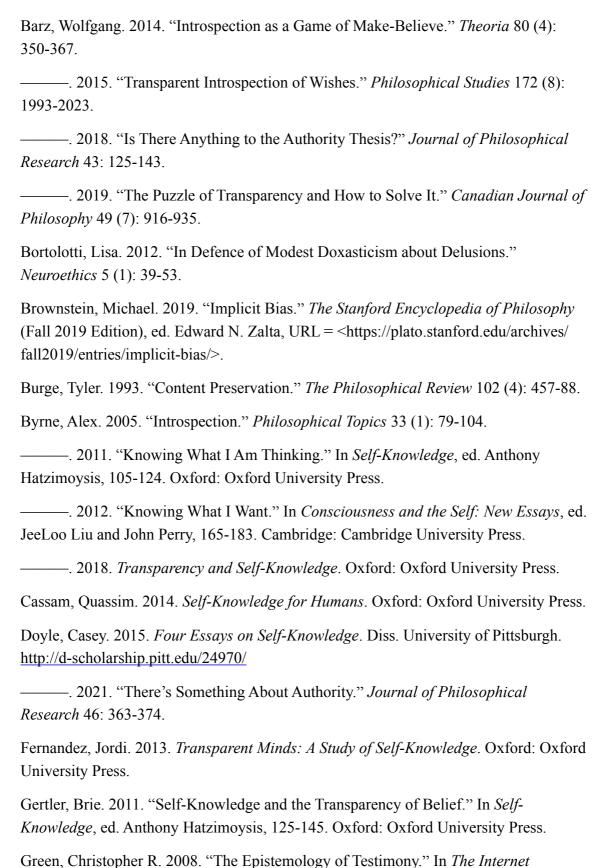
[9] See Barz 2014; 2015; 2019. For similar views, see Byrne 2005; 2011; 2012; 2018; and Fernandez 2013.

REFERENCES

Barz, Wolfgang. 2014. "Introspection as a Game of Make-Believe." *Theoria* 80 (4): 350-367.

———. 2015. "Transparent Introspection of Wishes." *Philosophical Studies* 172 (8): 1993-2023.

———. 2018. "Is There Anything to the Authority Thesis?" *Journal of Philosophical Research* 43: 125-143.

———. 2019. "The Puzzle of Transparency and How to Solve It." *Canadian Journal of Philosophy* 49 (7): 916-935.

Bortolotti, Lisa. 2012. "In Defence of Modest Doxasticism about Delusions." *Neuroethics* 5 (1): 39-53.

Brownstein, Michael. 2019. "Implicit Bias." *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), ed. Edward N. Zalta, URL = <https://plato.stanford.edu/archives/fall2019/entries/implicit-bias/>.

Burge, Tyler. 1993. "Content Preservation." *The Philosophical Review* 102 (4): 457-88.

Byrne, Alex. 2005. "Introspection." *Philosophical Topics* 33 (1): 79-104.

———. 2011. "Knowing What I Am Thinking." In *Self-Knowledge*, ed. Anthony Hatzimoysis, 105-124. Oxford: Oxford University Press.

———. 2012. "Knowing What I Want." In *Consciousness and the Self: New Essays*, ed. JeeLoo Liu and John Perry, 165-183. Cambridge: Cambridge University Press.

———. 2018. *Transparency and Self-Knowledge*. Oxford: Oxford University Press.

Cassam, Quassim. 2014. *Self-Knowledge for Humans*. Oxford: Oxford University Press.

Doyle, Casey. 2015. *Four Essays on Self-Knowledge*. Diss. University of Pittsburgh. http://d-scholarship.pitt.edu/24970/

———. 2021. "There's Something About Authority." *Journal of Philosophical Research* 46: 363-374.

Fernandez, Jordi. 2013. *Transparent Minds: A Study of Self-Knowledge*. Oxford: Oxford University Press.

Gertler, Brie. 2011. "Self-Knowledge and the Transparency of Belief." In *Self-Knowledge*, ed. Anthony Hatzimoysis, 125-145. Oxford: Oxford University Press.

Green, Christopher R. 2008. "The Epistemology of Testimony." In *The Internet Encyclopedia of Philosophy*, accessed August 01, 2023, https://iep.utm.edu/ep-testi/

Lackey, Jennifer. 2008. *Learning from Words: Testimony as a Source of Knowledge*. Oxford: Oxford University Press.

Lawlor, Krista. 2009. "Knowing What One Wants." *Philosophy and Phenomenological Research* 79 (1): 47-75.

Schwitzgebel, Eric. 2012. "Self-Ignorance." In *Consciousness and the Self: New Essays*, ed. JeeLoo Liu and John Perry, 184-197. Cambridge: Cambridge University Press.

Thompson, Michael. 2012. "Anscombe's Intention and Practical Knowledge." In *Essays on Anscombe's Intention*, ed. Anton Ford, Jennifer Hornsby, and Frederick Stoutland, 198-210. Cambridge, MA: Harvard University Press.

Winokur, Benjamin. 2022. "There Is Something to the Authority Thesis." *Journal of Philosophical Research* 47: 115-132.

———. 2023. "Authority as (qualified) indubitability." *Inquiry*, online first. DOI: 10.1080/0020174X.2023.2224543